STATISTICAL RESEARCH REPORT
Institute of Mathematics
University of Oslo

No 2
May 1974

SOME MONTE CARLO RESULTS CONCERNING THE PROBABILITIES OF FALSE REJECTIONS FOR A MULTIPLE HYPOTHESES TESTING PROBLEM

by

Tore Skrøppa

SOME MONTE CARLO RESULTS CONCERNING THE PROBABILITIES OF FALSE REJECTIONS FOR A MULTIPLE HYPOTHESES TESTING PROBLEM

In a multiple hypotheses testing problem where M individual hypotheses are tested at levels $\alpha_1,\dots,\alpha_M$, respectively, only an upper bound $\Sigma_{t=1}^{M}\alpha_t$ is generally known for the probability of at least one false rejection. Neither do we know the probabilities of exactly one false rejection, two false rejections, etc. In order to investigate these probabilities a Monte Carlo study has been performed.

The model assumed is:

$$X_{ij}\ ,\ j = 1,\dots,n\ ,\ i = 1,\dots,k$$

are independently, normally distributed random variables with a common known variance σ^2 and expectations $EY_{ij} = \mu_i$, $j = 1,\dots,n$, $i = 1,\dots,k$.

We will be concerned only with all possible pairwise comparisons, and hence, the multiple problem is

$$(1)\qquad H_{ij} : \mu_i = \mu_j \quad \text{against} \quad K_{ij} : \mu_i \neq \mu_j\ ,\ i < j\ .$$

Thus, the number of hypotheses tested equals $M = k(k-1)/2$.

A test of the problem (1) consists of rejecting H_{ij} and accepting K_{ij} when

$$(2)\qquad |\bar{X}_{i.} - \bar{X}_{j.}| > z_\rho (2/n)^{1/2}\ ,$$

where $\rho = \alpha/2M$ and z_ρ is the upper 100ρ percentage point of the standard normal distribution.

With this procedure the probability of at least one false rejection when all hypotheses are true, is less or equal to α . Denote this probability by P , and let P_t be the probability of exactly t false rejections when $\mu_1 = \dots = \mu_k$. We then have $P \leq \alpha$ and $P = P_1 + \dots + P_M$.

A series of random experiments were generated on the CDC 3300 computer using a well-tested procedure in Simula 67 which generates "random normal variaties" with known expectations and variances. For $k = 3,4,\dots,10$ $N = 10\ 000$ experiments were performed, each consisting of k normally distributed random variables with expectations 0 and variances $1/n$. In each experiment the hypothesis H_{ij} against the alternative K_{ij} was tested by (2), $i < j$, and the number of rejections was recorded. In all experiments $n = 10$ and $\alpha = 0.10$.

Let F_t be the random variable which tells the number of experiments with exactly t false rejections, $t = 1,\dots,M$. The probabilities P and P_t, $t = 1,\dots,M$ are then estimated by

(3) $$\hat{P}_t = F_t/N \ , \ t = 1,\dots,M$$

(4) $$\hat{P} = (F_1+\dots+F_M)/N$$

The standard errors of the estimates are given by $(\hat{P}_t(1-\hat{P}_t)/N)^{1/2}$, $t = 1,\dots,M$ and $(\hat{P}(1-\hat{P})/N)^{1/2}$, respectively.

The results of the experiments are given in Table 1, and Table 2 shows the highest estimate of the standard errors of $\hat{P}$ and $\hat{P}_t$, $t = 1,\dots,M$.

For all values of k the estimate of P is less than the upper bound 0.10, and the estimates are decreasing as k increases. This tendency is the one to be expected since the levels of the individual tests then become smaller and greater differences in the averages are necessary in order to achieve significance.

If the experimenter of these experiments was satisfied with $P=0.10$, then α could be increased to 0.12 or perhaps 0.13 - 0.14 for high values of k. This is due to the fact that the upper bound for P, α, is not achieved. The individual tests would then become more powerful.

Table 1 also shows that P_1 and P_2 are greater than $P_3, P_4, \ldots$. In experiments with false rejections ca. 70% of the cases were made up of only one false rejection, and ca. 20% of the cases were made up of two false rejections. No experiment with more than 6 false rejections was recorded.

It seems reasonable to believe that similar results hold for multiple hypotheses testing problems with other types of distributions and different types of dependency among the test statistics.

Table 1. Estimates of P and P_t, $t = 1,\ldots,M$.

k	P	P_1	P_2	P_3	P_4	P_5	P_6
3	.0793	.0648	.0145	0	-	-	-
4	.0809	.0615	.0159	.0032	.0003	0	0
5	.0760	.0569	.0133	.0047	.0010	.0001	0
6	.0709	.0507	.0142	.0042	.0015	.0002	.0001
7	.0739	.0523	.0160	.0038	.0014	.0004	0
8	.0726	.0503	.0157	.0048	.0014	.0003	.0001
9	.0680	.0462	.0151	.0042	.0018	.0006	.0001
10	.0680	.0468	.0151	.0043	.0011	.0005	.0002

Table 2. Highest standard errors of the estimates.

P	P_1	P_2	P_3	P_4	P_5	P_6
.0027	.0025	.0013	.0007	.0004	.0003	.0001